AFRL-AFOSR-VA-TR-2015-0330

SOCIOSCAPE: REAL-TIME ANALYSIS OF DYNAMIC HETEROGENEOUS NETWORKS IN COMPLEX SOCIO-CULTURAL SYSTEMS

Eric Xing
CARNEGIE MELLON UNIVERSITY
5000 FORBES AVENUE
PITTSBURGH, PA 15213-3815

10/22/2015
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/RTA2

Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
09-28-2015

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
06/01/2010 - 05/31/2015

4. TITLE AND SUBTITLE
SOCIOSCAPE: Real-Time Analysis of Dynamic Heterogeneous Networks in Complex Socio-Cultural Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA95501010247

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Xing, Poe Eric

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF, AFRL DUNS 143574726
AF OFFICE OF SCIENTIFIC RESEARCH
875 NORTH RANDOLPH STREET
ARLINGTON VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
N/A

12. DISTRIBUTION/AVAILABILITY STATEMENT
This report is approved for public release.
DISTRIBUTION A

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT
In many problems arising in social, technological, and other fields, it is often necessary to analyze populations of individuals interconnected by a network. Real-time analysis of network data is important for detecting anomaly, predicting vulnerability, and assessing the potential impact of interventions in various social and information systems. It is not unusual for network data to be large, dynamic, heterogeneous, noisy and incomplete. Each of these characteristics adds a degree of complexity to the interpretation and analysis of networks.
Traditional approaches to network analysis tend to make simplistic assumptions, such as assuming that there is only a single node or edge type, or ignoring the role/mind of nodal actors and the dynamics of the networks. We intend to develop new hierarchical and dynamic Bayesian formalisms and novel graph.

15. SUBJECT TERMS
network, time-varying network, machine learning, community detection, information extraction, inference, big data

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	SAR		Eric Poe Xing
					19b. TELEPHONE NUMBER *(Include area code)* 412 268 2559

Reset

DISTRIBUTION A: Distribution approved for public release.

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Real-time Analysis and Mining of Dynamic Networks in Complex SocioCultural Systems

Our proposed research consists of 5 research themes:

1: Develop new methods for latent theme distillation and data integration for network data; apply them to unravel the sociocultural driving forces behind network formation by reasoning the roles or activities of social actors, and the semantic or ideological underpinnings of social relationships in the network.
2: Develop new hierarchical formalisms for modeling and identifying sociocultural evolution or revolution underlying community coalescence or fragmentation; develop new dynamic nonparametric Bayesian formalisms for evolutionary clustering of network entities in open possible worlds, that is, worlds allowing unknown and unbounded number of communities with stochastic birth/death/transformation over time.
3: Develop new formalisms for modeling network rewiring over time; use them to investigate the mechanisms of social network dynamics under sociocultural variability.
4: Develop both model-based and optimization-based algorithms for the yet unexplored problem of reverse engineering unobserved temporally rewiring networks from time series of entity attributes.
5: Develop new theories and new algorithms for identifying and tracking information diffusion, community formation, and their impacts on networks.

Research:

Overall, we feel all the goals we set forth above have been successfully accomplished, and we exceeded the original objectives by also addressing the problem of scalable inference on massive networks of societal scale. Bellow, we briefly summarize major highlights in each direction, and the relevant papers. (The ordering of the themes are slightly modified to allow a more intuitive logic.)

Theme 1: network tomography on role prediction and evolution

We developed a new family of hierarchical and dynamic Bayesian latent space models for graphs and network data, which provide new theoretical frameworks for inferring latent semantic/functional aspects of graphs, and for modeling, interpreting and predicting graphs that evolve over time (such as dynamically re-wired biological networks and social networks).

In particular we developed the mixed membership of stochastic block model (MMSB) and GMF-based algorithms that can infer the hidden multi-involvement of each nodal actor in different roles and predict hidden links between nodes. We have developed a number of significant extensions of the MMSB model, including the mixed membership triangular model (MMTM) for inference latent role based on more informative network motif features, the joint network-text models for latent role and community inference, and the dynamic MMSB model for inferring trajectories of actors' states in latent role-space.

And recently, we have further developed now techniques for scalable mixed membership modeling that allow network with hundreds of millions nodes to be analyzed. This is an area that the research community has just started addressing.

E. Airodi, D. Blei, S. Fienberg and E. P. Xing, Mixed Membership Stochastic Blockmodels. Journal of Machine Learning Research, Vol 9:1981–2014, 2008.

E.P. Xing, W. Fu, and L. Song, A State-Space Mixed Membership Blockmodel for Dynamic Network Tomography, Annals of Applied Statistics, Vol. 4, No. 2, 535 - 566, 2010.

Q. Ho, J. Eisenstein and E. P. Xing, Document Hierarchies from Text and Links, Proceedings of the International World Wide Web Conference (WWW 2012).

Q. Ho, J. Yin and E. P. Xing, On Triangular versus Edge Representations — Towards Scalable Modeling of Networks. Advances in Neural Information Processing Systems 26 (NIPS '12).

J. Yin, Q. Ho and E. P. Xing, A Scalable Approach to Probabilistic Latent Space Inference of Large- Scale Networks. Advances in Neural Information Processing Systems 27 (NIPS ’13).

M. Sachan, A. Dubey, S. Srivastava, E. P. Xing and Eduard Hovy, Spatial Compactness meets Topical Consistency: Jointly modeling Links and Content for Community Detection , Proceedings of The 7th ACM International Conference on Web Search and Data Mining (WSDM 2014).

Theme 2: Dynamic information clustering and group tracking,

We have extends our model for dynamic network evolution to allow inference of multimodel dynamic trajectories of actor roles over time. We found that this new model fit real life data significantly better and enable inference of rich and more realistic role evolution. This model represents each actor in the network by its latent function(s) in a simplex, and tracks the evolution of these latent functions across time. It features a clustering, temporal logistic normal model, which:

1) Captures interesting covariance structures between latent functions
2) Facilitates temporal modeling of latent functions
3) Clusters similar latent function configurations for better a statistical fit on multi-modal data

We call our model dM3SB or Dynamic Mixture of Mixed Membership Stochastic Blockmodels, and have developed efficient approximate inference and learning algorithms for it. When applied to social networks such as the US senator network, dM3SB offers extremely interesting insight into the network community structure and actor behavior, such as both parties interact exclusively with themselves, and how certain moderate or swing senators networking with ones from opposite parties.

News clustering, categorization and analysis are key components of any news portal. They require algorithms capable of dealing with dynamic data to cluster, interpret and to temporally aggregate news articles. These three tasks are often solved separately. We have developed a unified framework to group incoming news articles into temporary but tightly-focused storylines, to identify prevalent topics and key entities within these stories, and to reveal the temporal structure of stories as they evolve. We achieve this by building a hybrid clustering and topic model. To deal with the available wealth of data we build an efficient parallel inference algorithm by sequential Monte Carlo estimation. Time and memory costs are nearly constant in the length of the history, and the approach scales to hundreds of thousands of documents. We demonstrate the efficiency and accuracy on the publicly available TDT dataset and data of a major internet news site with performance that compares favorably to the state of the art.

Q. Ho, L. Song and E. P. Xing, Evolving Cluster Mixed-Membership Blockmodel for Time-Evolving Networks, Proceedings of the 14th International Conference on Artifical Intelligence and Statistics (AISTAT 2011).

A. Ahmed and E. P. Xing, Timeline: A Dynamic Hierarchical Dirichlet Process Model for Recovering Birth/Death and Evolution of Topics in Text Stream, Proceedings of the 26th International Conference on Conference on Uncertainty in Artificial Intelli- gence (UAI 2010).

Amr Ahmed, Qirong Ho, Choon-hui Teo, Jacob Eisenstein, Alex Somla, Eric P. Xing. Online Inference for the Infinite Cluster-topic Model: Storylines from Streaming Text. AISTATS 2011.

Amr Ahmed, Qirong Ho, Jacob Eisenstein, Eric P. Xing, Alex Somla, Choon-Hui Teo. Unified Analysis of Streaming News. WWW 2011.

Theme 3: Relationship prediction and evolution,

We have worked on Link prediction from Text and Network Features. Social media services such as Twitter make such social connections explicit, and our research has explored how to model this relationship more directly. One intriguing result from this work is that text can predict hidden social connections quite effectively; indeed, using a topic model of text (while ignoring network structure), we are able to obtain predictions that outperform a network baseline. We trained a topic model on 21,000 users of Twitter, and predicted links based on topical similarity. The table above shows results for predicting which users will send each other messages (left) and which users will

follow each other (right). The X-axis shows the results for different number of topics, and the color indicates whether post-hoc regression was applied to tune the importance of each topic towards the similarity metric. The Y-axis shows the average rank of predicted links; lower scores is better, and chance is 10,500. In all cases, topic-based link prediction strongly outperforms a link prediction heuristic that counts the number of shared neighbors between two nodes. We are now developing joint models that incorporate both text and network features.

We have developed the theory of non-degeneracy of temporal exponential random graph model over social networks evolution, proving that it's maximum likelihood estimator is always non-degenerate, unlike the sometimes degenerate conventional exponential random graph models. We showed that, using this model, one can perform hypothesis tests over multiple different link/motif evolution dynamics, predicting actor label, and simulate more realistic evolving social networks based on different patterns of social interactions, rather than simply scale-free graphs.

Papers relevant to this theme include:

Kriti Puniyani, Jacob Eisenstein, Shay Cohen and Eric P. Xing. Social Links from Latent Topics in Microblogs. Proceedings of the NAACL Workshop on Social Media, 2010. Winner of best presentation award.

S. Hanneke, W. Fu and E. P. Xing, Discrete Temporal Models of Social Networks, Electronic Journal of Statistics Vol. 4 (2010) 585-605.

Theme 4: inferring unobservable changing networks

We have developed a family of nonparametric estimators of time-evolving or tree-evolving graphical models, including evolving Gaussian Graphical Models (GGM), Markov Random Fiends (MRF), and Auto-Regressive Dynamic Bayesian Networks, based on novel extensions of the graphical lasso technique originally used for sparsistent structure recovery of time-invariant GGMs and MRFs. The property of sparsistency we were able to prove for our estimators is an important characteristic of these type of estimator, because it reveals conditions where correct recovery of network structure under various models is possible even when the size of the graph is very large (i.e., tens of thousands of nodes) whereas the number of samples of nodal state are small (i.e, 101 ~ 102). Our estimator includes TESLA (based on temporally-smoothed and regularized graphical regression), KELLER (based on kernel reweighted regularized graphical regression), and a number of other fancier versions. We have successfully used them for reverse-engineering latent evolving social networks in the US Senate and the Enron corporation, the evolving gene network of fruit fly while aging, and the gene networks evolving along cell lineage during breast cancer progression and reversal, at a time resolution only limited by sample frequency.

M. Kolar, and E. P. Xing, Estimating Time-Varying Networks With Jumps. Electronic Journal of Statistics Vol. 6 (2012) 2069-2106.

M. Kolar, H. Liu and E. P. Xing, Graph Estimation From Multi-attribute Data. Journal of Machine Learning Research, in press, 2014.

M. Kolar and E. P. Xing, Ultra-high Dimensional Multiple Output Learning With Simultaneous Orthogonal Matching Pursuit, Proceedings of the 13th International Conference on Artifical Intelligence and Statistics (AISTAT 2010).

M. Kolar, L. Song, A. Ahmed, and E. P. Xing, Estimating Time-Varying Networks. Annals of Applied Statistics, Vol. 4, No. 1, 94 – 123, 2010 (arXiv:0812.5087).

M. Kolar, L. Song, A. Ahmed, and E. P. Xing, Estimating Time-Varying Networks. Annals of Applied Statistics, Vol. 4, No. 1, 94123, 2010.

A. Ahmed and E. P. Xing, Recovering Time-Varying Networks of Dependencies in Social and Biological Studies. Proc. Natl. Acad. Sci., vol. 106, no. 29, 11878-11883, 2009.

L. Song, M. Kolar and E. P. Xing, KELLER: Estimating Time-Evolving Interactions Between Genes, Bioinformatics 2009 25(12): i128-i136. (Proceedings of ISMB '09)

M. Kolar and E. P. Xing, Sparsistent Learning of Varying-coefficient Models with Structural Changes. Advances in Neural Information Processing Systems 23, MIT Press, Cambridge, MA, 2010. (NIPS '09).

L. Song, M. Kolar and E. P. Xing, Time-Varying Dynamic Bayesian Networks. Advances in Neural Information Processing Systems 23, MIT Press, Cambridge, MA, 2010. (NIPS '09).

Theme 5: Geographic Variation of linguistic communities, and Predicting Author Demographics from Social Media Text

Decades of sociolinguistic research have documented the strong and complex relationship between language and the demographic components of personal identity, such as race, class and gender. However, such research has relied on the intuition of the investigator to identify the relevant linguistic indicators of demographic membership, an approach which cannot easily be applied in prediction scenarios and new domains.

Borrowing techniques from genome analysis, we have developed a method for identifying sociolinguistic associations from social media text and widely-available metadata. While there are thousands or millions of potential associations between linguistic features and demographic attributes, we apply composite sparsity-inducing regularizers to induce a small dictionary of linguistic features with strong demographic associations. This enables the first accurate predictions of the race, ethnicity, and other demographic features from raw text alone.

Moreover, by inducing models with structured sparsity, our approach facilitates new sociolinguistic insights, such as significant racial differences in the usage of internet idioms such as emoticons (largely used by whites) and abbreviations (e.g., smh / "shake my head", largely used by minorities).

Through statistical inference, we recover a model of the relationship between text and geographical communities. We test the fidelity of this model by attempting to predict the geographical location of authors from their text alone. Our median error is 500 kilometers (less than the distance from Los Angeles to San Francisco); we predict the correct state 28% of the time. These results compare very favorably with alternative approaches that do not construct geographical communities but instead relate text directly to geographical coordinates.

J. Eisenstein, B. O'Connor, N. A. Smith, and E. P. Xing, A Latent Variable Model for Geographic Lexical Variation, 2010 Conference on Empirical Methods on Natural Language Processing (EMNLP 2010).

Jacob Eisenstein, Noah A. Smith, and Eric P. Xing. Discovering Sociolinguistic Associations with Structured Sparsity. Proceedings of ACL 2011.

Jacob Eisenstein, Amr Ahmed, and Eric P. Xing. Sparse Additive Generative Models of Text. Proceedings of ICML 2011.

Additional Work:

In addition to the above-proposed work, we have also investigated a number of other directions closely related to the problems address in this project.

Towards societal-scale analysis of social networks

In the context of network analysis, a latent space refers to a space of unobserved latent representations of individual entities (i.e., topics, roles, or simply embeddings, depending on how users would interpret them) that govern the potential patterns of network relations. The problem of latent space inference amounts to learning the bases of such a space and reducing the high-dimensional network data to such a lower-dimensional space, in which each entity has a position vector. Depending on model semantics, the position vectors can be used for diverse tasks such as community detection, user personalization, link prediction and exploratory analysis. However, scalability is a key challenge for many existing probabilistic methods, as even recent state-of-the-art methods still require days to process modest networks of around 100,000 nodes.

We have developed a scalable approach, called the Parsimonious Triangular Model (PTM), for making inference about latent spaces of large networks. With a succinct representation of networks as a bag of triangular motifs, a parsimonious statistical model, and an efficient stochastic variational inference algorithm, we are able to analyze real networks with over a million vertices and hundreds of latent roles on a single machine in a matter of hours, a setting that is out of reach for many existing methods. When compared to the state-of-the-art probabilistic approaches, our method is several orders of magnitude faster, with competitive or improved accuracy for latent space recovery and link prediction.

When compared to the Mixed Membership Stochastic Blockmodel (a popular network latent space model), our PTM not only scales to much bigger networks (in excess of 1 million nodes), but also reports competitive or even improved link prediction results, as shown in the table below.

PTM not only scales to large networks; it also completes inference on them in a matter of hours. For example, our PTM inference algorithm took only the 1.1 million-node Youtube network with 100 latent roles took only 3h to converge on a single multicore machine,

J. Yin, Q. Ho and E. P. Xing, A Scalable Approach to Probabilistic Latent Space Inference of Large-Scale Networks. Neural Information Processing Systems, 2013 (NIPS 2013).

Other Activities:

I have given several research and tutorial talk on network sciences and related statistical methodologies:

[1] Dynamic Network Analysis: Model, Algorithm, Theory, and Application,
Columbia Statistics Seminar, Columbia University, New York, October 11, 2010.

[2] Reverse Engineering Tree-Evolving Gene Networks Underlying Developing Breast Cancer Cell Lineages, Stanford CCSB Seminar, Center for Cancer Systems Biology, Stanford University, Palo Alto, CA, November 20, 2010.

[3] Learning varying coefficient varying structure models: Reverse engineering rewiring networks underlying dynamics processes, Stanford Statistics Seminar, Department of Statistics, Stanford University, Palo Alto, CA, January 18, 2011.

[4] Probabilistic Graphical Models: Theory, Algorithms and Application, Compact Course, UniversitŁt Heidelberg, Germany, February 7-11, 2011.

[5] On High-Dimensional Sparse Structured Input-Output Models, with Applications to Genome-Phenome Association Analysis of Complex Diseases, Workshop in Biostatistics, Department of Statistics, Stanford University, Palo Alto, CA, February 24, 2011.

[6] Topic Models, Latent Space Models, Sparse Coding, and All That: A systematic understanding of probabilistic semantic extraction in large corpus, Tutorial: The 50th Annual Meeting of the Association for Computational Linguistics, (ACL 2012), Jeju , Korea, July 8-11, 2012.

[7] Machine Learning Approaches to Network and Social Media, Distinguished Lecture Series, George Mason University, Washington DC, April 19, 2013.

[8] Big Data, Big Model, and Big Learning, CS Distinguished Lecture, University of Southern California, Los Angeles, May 22, 2013.

Student/Postdoc Training:

This project creates a platform that allows the following students and postdocs to be trained:

Current:

Kumar Avinava Dubey

Graduated:

Qirong Ho (now adjust assistant professor, Singapore Management University, KDD 2015 best dissertation runner up)
Steve Hanneke (now Asst. Prof. stat@CMU)
Wenjie Fu (now Software Engineer at Facebook)
Amr Ahmed (now Research Scientist at Google, KDD 2012 best dissertation)
Mladen Kolar (now Assistant Professor at U. of Chicago)
Le Song (Asst. Prof. cs@ Georgia Tech)
Jacob Eisenstein (Asst Prof. cs@ Georgia Tech)

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

epxing@cs.cmu.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

412-268-2559

Organization / Institution name

Carnegie Mellon University

Grant/Contract Title
The full title of the funded effort.

SocioScape: Real-time Analysis and Mining of Dynamic
Heterogeneous Networks in Complex Socio-Cultural Systems

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-10-1-0247

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

FA95501010247

Program Manager
The AFOSR Program Manager currently assigned to the award

Dr. Robert Bonneau

Reporting Period Start Date

06/01/2010

Reporting Period End Date

05/31/2015

Abstract

In many problems arising in social, technological, and other fields, it is often necessary to analyze populations
of individuals interconnected by a network. Real-time analysis of network data is important for detecting anomaly, predicting vulnerability, and assessing the potential impact of interventions in various social and information systems. It is not unusual for network data to be large, dynamic, heterogeneous, noisy and incomplete. Each of these characteristics adds a degree of complexity to the interpretation and analysis of networks.
Traditional approaches to network analysis tend to make simplistic assumptions, such as assuming that there is only a single node or edge type, or ignoring the role/mind of nodal actors and the dynamics of the networks. We intend to develop new hierarchical and dynamic Bayesian formalisms and novel graph evolution models for analyzing dynamic heterogeneous networks.
Our approach will build on the most recent advances in machine learning and statistical network analysis toward rich, multi-faceted network representations, and the most recent advances in stochastic process-based
approaches which incorporate rich dynamics. We will focus on answering useful analytic queries, such as,

hidden identity/role induction, structural/organizational forecast, system robustness, etc., particularly in the context of understanding culturally determined behavior of large groups and communities over time. Many of
such queries are relevant to applications important to national interests. For example, identifying asymmetric
threat based on limited observations hidden in volumes of complex, heterogeneous network.
Having such a unified framework will both help to advance the theory and methodology for understanding/ predicting networks, by providing a useful toolkit and an generic paradigm for computational inference and learning; and will also ensure that our methods are useful to analysts and extendable for tasks to be identified in the future. To further this goal, we plan to validate our methods on concrete socio-cultural and cyber domains.

Distribution Statement

This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

AFD-070820-035-xing-afosr[1].pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

AFOSR-report15.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

Changes in research objectives (if any):

Change in AFOSR Program Manager, if any:

Extensions granted or milestones slipped, if any:

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

Sep 28, 2015 14:50:45 Success: Email Sent to: epxing@cs.cmu.edu